Astrophysics, Computer Science, and Vedic Cosmologies

A Critical and Computational Study of Vedic and Upanishadic Ideas in Relation to Modern Astrophysics

Deekshith Komira

Independent Research Paper

Interdisciplinary: Astrophysics • Computer Science • History of Science • Sanskrit Studies

Date: December 19, 2025

Abstract

Public discussions sometimes claim that the Vedas and Upanishads “contain” modern astrophysics. Such claims usually fail because they collapse distinct genres and methods: poetic or metaphysical language is treated as if it were quantitative, testable physics. This paper proposes a more rigorous alternative. It reads Vedic and Upanishadic materials as sources for early South Asian cosmological imagination and for the cultural centrality of timekeeping; it traces the historical bridge from ritual calendars to explicit astronomical computation, especially through Vedanga Jyotisha and the later siddhanta tradition; and it shows how modern computer science—algorithm design, data modeling, NLP, and digital humanities workflows—can enable reproducible analysis of Sanskrit astronomical corpora and computational reconstruction of calendar procedures.

The paper distinguishes astrophysics and modern cosmology from premodern cosmology as metaphysical discourse. It treats textual motifs (origin, ṛta, cyclic time, and ākāśa) as comparative philosophy rather than scientific equivalence. It concludes with an implementable research plan: digitize Sanskrit astronomical texts from open repositories; apply Sanskrit segmentation and morphology; represent calendrical concepts (tithi, nakshatra, solar ingress) in a knowledge graph; implement rule engines; and validate computed calendars against reference ephemerides.

Keywords: Vedas; Upanishads; Vedanga Jyotisha; Indian astronomy; cosmology; calendrical calculation; Sanskrit NLP; digital humanities; computational philology; knowledge graphs; astrophysics.

Table of Contents

1. Introduction

2. Definitions, Scope, and Methodological Ground Rules

3. Vedic and Upanishadic Cosmology as Textual Evidence

4. From Ritual Timekeeping to Early Astronomy: Vedanga Jyotisha

5. Post-Vedic Mathematical Astronomy: The Siddhanta Trajectory

6. Comparative Cosmology: Productive Parallels and Category Errors

7. Computer Science Lens: Calendrical Algorithms and Computational Astronomy

8. Digital Humanities and NLP for Sanskrit Astronomical Corpora

9. Research Design: A Reproducible Computational Project

10. Ethics, Epistemics, and Communication

11. Conclusion

References

Appendices

List of Figures

- Figure 1. High-level timeline (schematic; dates approximate).
- Figure 2. Layered method (schematic).
- Figure 3. Planck spectrum at 2.725 K (CMB-like blackbody; schematic).
- Figure 4. Synthetic light curve (illustrative).
- Figure 5. Text-to-analysis pipeline (schematic).

1. Introduction

Figure 1 provides context for the historical layers and modern scientific domains discussed.

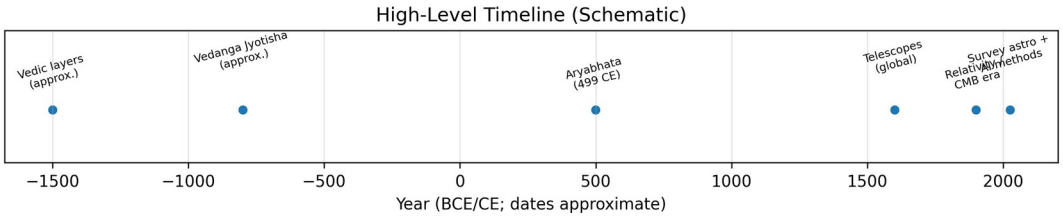

Figure 1. High-level timeline (schematic; dates approximate).

This paper addresses an interdisciplinary question that is widely asked but rarely answered with appropriate rigor: How are the Vedas and Upanishads “used” in astrophysics? In modern scientific practice, the answer is straightforward: they are not used as observational evidence, data sources, or physical theory. Astrophysics relies on telescopes and detectors, calibration and error models, statistical inference, and mathematical models grounded in physics. A scientific claim must, at minimum, make predictions that can be tested against observation.

However, the question remains valuable if it is reframed. The Vedas and Upanishads can be “used” in adjacent scholarly and technical ways: as primary sources for the history of cosmological thought; as cultural contexts that motivated precise calendrical timekeeping; and as

corpora for computational analysis using modern methods (digitization, OCR, NLP, and knowledge graphs). When the question is framed as a research program, it becomes both intellectually serious and practically implementable.

I, Deekshith Komira, write from a personal intersection of training and interest: I have studied Vedic literature in a traditional setting, while also pursuing topics in astrophysics, cosmology, mythology, and computer science. This positionality makes it especially important to separate devotional significance from scientific claims and to respect both textual genre and scientific method. The aim is neither apologetic nor dismissive: it is to produce a credible, research-grade synthesis.

1.1 Research questions and contributions

This paper addresses four research questions:

- What cosmological and temporal motifs appear in Vedic and Upanishadic literature, and what are their genres and purposes?
- What computational and observational content is present in early calendrical astronomy associated with Vedic practice, especially Vedanga Jyotisha?
- How did later Sanskrit astronomical traditions (siddhantas) operationalize time and celestial motion, and how should they be related to earlier Vedic materials?
- What computer science methods can enable a reproducible analysis of Sanskrit astronomical texts and their calendrical computations?

The contributions are: (i) a methodological framework for evaluating “Vedas and astrophysics” claims; (ii) a historically grounded bridge from textual cosmology to calendrical astronomy; (iii) a computer-science-centered view of calendrical algorithms; and (iv) a concrete implementation plan for a computational research project.

1.2 Scope limitations

Two limitations are deliberate. First, this is not a theological argument about truth claims of scripture; it is a research paper about texts, computation, and scientific method. Second, it does not attempt to “prove” that any verse predicts a specific modern cosmological parameter. Instead, it aims to build a defensible bridge: conceptual motivation in early texts, procedural computation in later astronomy, and modern computational methods for analyzing that historical record.

2. Definitions, Scope, and Methodological Ground Rules

The paper separates interpretive parallels from physical equivalences using a layered method.

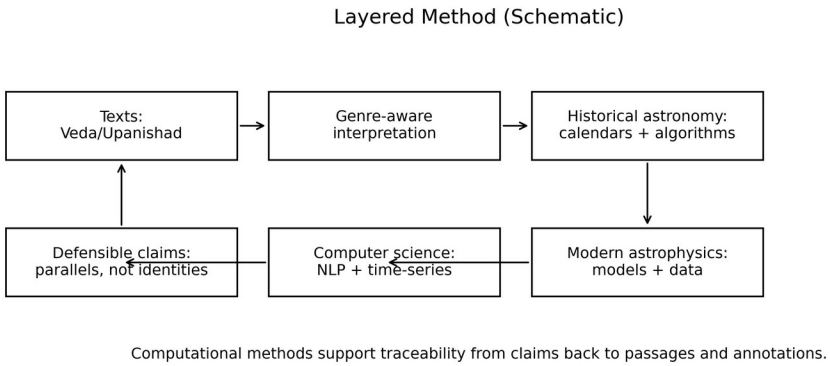

Figure 2. Layered method (schematic).

2.1 Astronomy, astrophysics, and cosmology

Astronomy is the study of celestial objects and phenomena. Astrophysics applies physical theory and quantitative models to explain astronomical observations. Cosmology studies the Universe as a whole: its origin, content, large-scale structure, and evolution. Modern cosmology is strongly supported by multiple lines of evidence, including measurements of the cosmic microwave background (CMB), interpreted as relic radiation from the early Universe, and

precision surveys of galaxies and supernovae. These evidence streams constrain models such as the standard ΛCDM framework.

In this paper, ‘astrophysics’ is used in its standard scientific meaning: it is not a synonym for ‘cosmic philosophy’ or ‘mythology.’ When we compare premodern cosmological statements to modern science, the comparison is clearly labeled as philosophical or historical rather than scientific equivalence.

2.2 Textual criticism and genre awareness

Vedic hymns are liturgical-poetic; Upanishads are philosophical and dialogical; and later astronomical treatises are technical and procedural. These genres cannot be treated as interchangeable. A hymn can be profound without being a measurement report. A philosophical dialogue can use cosmological imagery as a vehicle for metaphysics rather than for physical modeling. A technical verse can specify a procedure without justifying it in terms of underlying dynamics.

2.3 Layered methodology

To avoid anachronism, the paper separates three layers:

- Layer A: Textual cosmology (Vedas/Upanishads) as primary sources for ideas of origin, order, time, and space.
- Layer B: Historical astronomy (Vedanga Jyotisha, siddhantas) as sources for procedural computation and observational anchoring.
- Layer C: Modern astrophysics/cosmology as empirically validated scientific models and datasets.

The paper connects Layer A to Layer B historically and culturally, and compares Layer A to Layer C only in a limited philosophical sense. It does not claim that Vedic or Upanishadic passages are equivalent to modern cosmological theories.

2.4 Criteria for defensible claims

A claim in this paper is treated as defensible if it meets at least one of the following criteria:

- Textual: supported by close reading in genre context and consistent translation choices.
- Historical: consistent with the historiography of Indian astronomy and manuscript transmission.
- Computational: implementable as an algorithm and testable with reproducible code.
- Scientific: (rare here) supported by observational data and current physical theory.

This explicit criteria list functions as a guardrail. It prevents the most common failure mode in "ancient science" discussions: moving from a suggestive metaphor to an unwarranted scientific conclusion.

3. Vedic and Upanishadic Cosmology as Textual Evidence

3.1 Creation and uncertainty

The Nāsadīya Sūkta (Rig Veda 10.129) is often cited because it approaches creation as an open question. Its rhetoric is notable for epistemic humility—the suggestion that even divine beings may not know the origin. In this paper, the hymn is treated as evidence about intellectual attitudes toward origins, not as a quantitative cosmology. This distinction is important because it allows comparison to modern scientific norms (provisionality, uncertainty) without making a false claim about predictive physics.

3.2 Ṛta and lawful order

The concept of ṛta is frequently glossed as cosmic order or truth. In Vedic contexts it underwrites the expectation that the world—including the heavens—exhibits patterned regularities. While this is not physics, it helps explain why systematic calendrical synchronization became culturally central: it operationalizes cosmic regularity in communal practice. In a systems view, ṛta can be seen as a normative commitment to regularity that supports the long-term maintenance of calendars, observatories, and educational transmission.

3.3 Kāla: time as metaphysical and practical

Time (kāla) operates at multiple levels: metaphysical reflection, ritual scheduling, and seasonal life. The practical layer matters most for astronomy: ritual schedules, festivals, and agricultural rhythms create demand for stable time-reckoning systems. In modern terms, this is a user-requirement story: a calendar is a social technology built atop astronomical cycles.

A useful computational analogy is to distinguish ‘time as concept’ from ‘time as interface.’ Even if a community holds metaphysical views about time, it still needs an operational interface—a set of conventions and calculations that map celestial cycles to actionable dates.

3.4 Ākāśa and the metaphysics of space

Upanishadic discussions of ākāśa (often rendered ‘space’ or ‘ether’) belong to metaphysics, not to spacetime physics. But they are relevant to the intellectual history of spatial concepts and to comparative philosophy of science. A responsible comparison distinguishes metaphorical or metaphysical ‘subtle elements’ from modern physical fields. Where modern physics speaks of spacetime geometry and fields with measurable interactions, Upanishadic language is concerned with ontological categories and the conditions for experience and liberation.

3.5 Cyclic time and scale

Many Indian traditions develop cyclic models of cosmic time, later articulated as yugas and kalpas with immense scales. Modern cosmology also works on vast timescales, but similarity of magnitude is not methodological equivalence. Cyclic metaphysics can coexist with empirical astronomy; it does not, by itself, imply modern cosmic evolution models.

From a philosophy-of-science standpoint, the key difference is operationalization. Modern timescales are inferred from specific measurements and models (stellar evolution, expansion history, radioactive dating). Mythic or metaphysical timescales do not typically specify such inference chains. The research value is therefore interpretive and historical, not predictive.

3.6 Interim conclusion

The Vedas and Upanishads provide a conceptual vocabulary—origin, order, time, and space—rather than predictive astrophysical models. Their most plausible technical legacy is indirect: they support the cultural priority of synchronizing human action with celestial regularities, creating a context in which calendars and observational attention become essential.

4. From Ritual Timekeeping to Early Astronomy: Vedanga Jyotisha

Vedanga Jyotisha is commonly viewed as a bridge between Vedic ritual practice and explicit astronomical computation. Its practical aim is calendrical: coordinating lunar and solar cycles to determine ritual timing. In modern terms, it resembles a rule specification for a lunisolar calendar, using average motions and named observational anchors such as nakshatras and solstitial markers.

4.1 What it computes

The text formalizes time units, cycle counts, and relationships between solar and lunar measures. This is not astrophysics; it is calendrical engineering. The scientific content lies in procedural specification: how to compute dates, months, and periodicities. It represents an early instance of turning observational regularities into reusable computation rules.

4.2 Observational anchoring through nakshatras

The nakshatra system ties calendrical reasoning to repeated sky patterns by naming lunar stations along the ecliptic. As a computational object, nakshatras can be represented as an ontology (a finite ordered set with associated attributes and rules). This representation supports automated analysis of how nakshatra terminology changes across texts and epochs.

4.3 Historiography and transmission

The dating and contextual origins of early Indian mathematical astronomy are debated. Historians of science emphasize that astronomical systems evolve through textual transmission, observational practice, and cross-cultural exchange, rather than emerging fully formed from poetic hymns. This contextual framing is essential for any scholarly paper linking ‘Vedic’ materials to astronomy.

4.4 What “use” means here

In a research context, the Vedic connection is best described as adjacency and motivation: Vedic practice created strong calendrical demand; Vedanga Jyotisha operationalized that demand with rules; and later traditions refined those computations. Modern astrophysics does not use this as physical theory, but computer science can use it as a domain for algorithmic reconstruction and validation.

4.5 Computational note: reconstructing an epoch

A technically interesting problem is reconstructing the epoch implied by certain solstitial or nakshatra references. This requires careful historical and astronomical handling because transmitted texts can preserve older astronomical markers. A computational study would: (i) encode the claimed solstice position; (ii) use modern precession models to estimate candidate epochs; and (iii) report uncertainty intervals rather than a single date. This is a strong example of how modern astronomy can be applied to historical texts without claiming that the texts are modern astrophysics.

5. Post-Vedic Mathematical Astronomy: The Siddhanta Trajectory

To understand the computational sophistication of South Asian astronomy, it is necessary to examine the siddhanta tradition—technical Sanskrit treatises that specify planetary models, trigonometric methods, and predictive procedures. This is where the strongest premodern scientific content is found.

5.1 Aryabhata and kinematic explanation

Aryabhata is widely credited with major advances in mathematical astronomy. A key historical point is the explanation of apparent daily motion as arising from Earth's rotation—a kinematic hypothesis that reframes observation as relative motion. This is methodologically important: it shows a willingness to propose explanatory mechanisms rather than merely describing patterns.

5.2 Tables and interpolation as computational primitives

Indian astronomical computation used tables (e.g., sine tables) and interpolation schemes, enabling efficient manual calculation. From a computer science perspective, tables and

interpolation are early examples of data structures and approximation algorithms designed under severe resource constraints (no calculators, limited notation, and human compute).

A modern parallel is the use of precomputed lookup tables in embedded systems or on-device inference, where compute resources are limited. This analogy can be used pedagogically: historical astronomy shows that engineering constraints shape mathematical technique.

5.3 Surya Siddhanta as an algorithmic ephemeris

The Surya Siddhanta and related texts provide rule-based procedures for computing solar, lunar, and planetary positions in a geocentric framework, with parameters and corrections. Modern celestial mechanics has different foundations, but as an algorithmic system the siddhanta corpus can be implemented and evaluated empirically.

5.4 Varahamihira and synthesis

Varahamihira's compilatory works (such as the Panchasiddhantika) illustrate another pattern in scientific history: synthesis of multiple parameter sets and model traditions. In computational terms, this looks like maintaining multiple model backends with a common interface. A modern research project might implement multiple siddhanta parameterizations and compare their predictive behavior across epochs.

5.5 Where 'Vedic' fits in this trajectory

It is tempting to label the entire Sanskrit scientific tradition as 'Vedic,' but that collapses important distinctions. The Vedas and early Upanishads belong to a different period and genre than siddhanta astronomy. A rigorous paper therefore uses precise language: Vedic and Upanishadic texts provide cosmological vocabulary and calendrical motivation; Vedanga Jyotisha is an early computational bridge; siddhantas deliver mature mathematical astronomy.

5.6 Implications for a modern interdisciplinary paper

This trajectory clarifies what a modern interdisciplinary paper can legitimately claim. The strongest claim is not that hymns encode modern physics, but that a long-running intellectual tradition cultivated computation, regularity, and procedural modeling. Computer science makes those procedures newly accessible by enabling large-scale digitization, rule extraction, and algorithmic evaluation.

6. Comparative Cosmology: Productive Parallels and Category Errors

Modern cosmology is constrained by quantitative observables; the near-blackbody CMB spectrum is a canonical example.

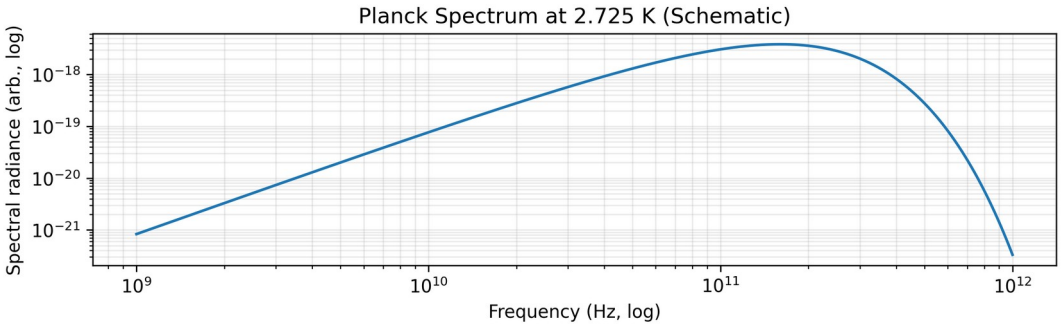

Figure 3. Planck spectrum at 2.725 K (CMB-like blackbody; schematic).

6.1 Productive parallels: inquiry and humility

The most defensible parallels between Vedic/Upanishadic cosmology and modern cosmology are philosophical: questions about origin, limits of knowledge, and the ordering of phenomena. These can be compared to modern scientific norms of provisionality and uncertainty quantification without claiming scientific identity.

6.2 Modern cosmology’s evidentiary anchors

Modern cosmology is grounded in precision measurements. The CMB is a central example: it is treated as relic radiation from the early Universe and provides a map of anisotropies that constrain cosmological parameters. Space missions such as WMAP and Planck illustrate how observational programs produce quantitative constraints and how public science agencies communicate these results.

The presence of such measurement programs highlights a key methodological difference: modern cosmology is inseparable from instrumentation and statistical inference. A premodern text can be historically important without playing the same evidentiary role.

6.3 Cycles vs. evolution

Cyclic metaphysics (creation–maintenance–dissolution) and evolutionary cosmology (hot Big Bang, structure formation) address different explanatory needs. A comparison can be made at the level of formal models, but a text-only analogy is insufficient. Scientific cyclic models, if proposed, must still be tested against datasets such as the CMB and large-scale structure.

6.4 Common category errors to avoid

- Treating metaphorical language (e.g., ‘ether’) as if it were a modern physical field.
- Equating spiritual ‘vibration’ or ‘sound’ metaphors with physical perturbations in early-Universe physics.
- Using large timescales in mythic cycles as ‘evidence’ of modern geologic or cosmological ages without a measurement method.
- Ignoring genre: hymns and dialogues are not procedural scientific specifications.

6.5 Interim conclusion

Comparisons can enrich philosophical reflection, but modern astrophysics remains methodologically distinct. The most substantial bridge between the traditions is computational: calendars and procedural astronomy.

7. Computer Science Lens: Calendrical Algorithms and Computational Astronomy

A reproducible pipeline supports evidence-linked claims and transparent cross-text comparison.

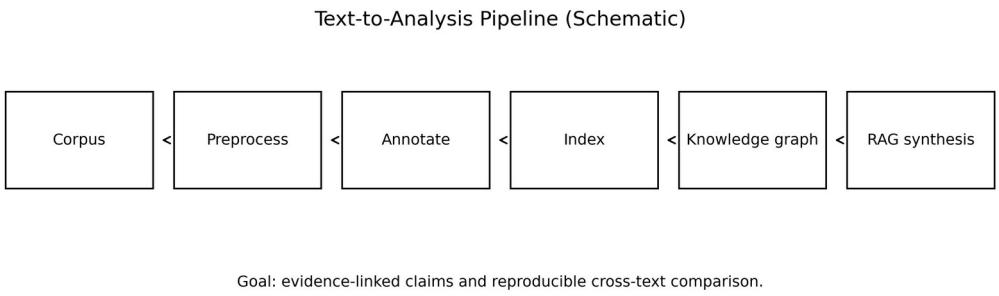

Figure 5. Text-to-analysis pipeline (schematic).

7.1 Calendars as algorithms

Calendars are algorithmic artifacts: they map continuous astronomical cycles into discrete dates and events. This makes them ideal for a paper at the intersection of computer science and historical astronomy. Algorithmic descriptions of Indian calendars have been developed in computer science contexts, emphasizing implementability and reproducibility.

7.2 Domain concepts: tithi, nakshatra, yoga, karana

Many panchang systems compute five daily elements (aṅgams):

- tithi: lunar day, based on Sun–Moon angular separation thresholds

- nakshatra: lunar mansion, based on the Moon’s ecliptic longitude interval
- yoga: derived from combined longitudes of Sun and Moon
- karana: half-tithi subdivisions used for finer timing
- vāra: weekday, linked to civil timekeeping

7.3 Event detection and numerical methods

Computationally, tithi and related elements are event-detection problems: find times at which an angular function crosses multiples of a threshold. This is solved by root-finding or bracketing methods under specified tolerances. The same numerical thinking appears in modern astrophysical data processing, where event detection and thresholding occur in transit searches, variable star detection, and signal processing.

7.4 Mean versus true motion and approximation tradeoffs

Premodern calendars often use mean motions (simple, stable) with correction terms. Modern implementations can use high-precision ephemerides, at greater compute cost. This tradeoff is directly legible to computer scientists as a balance among accuracy, performance, and interpretability.

7.5 Computational astrophysics as a parallel domain

Time-series analysis is a standard interface between computer science and astrophysics (e.g., variability and transit detection).

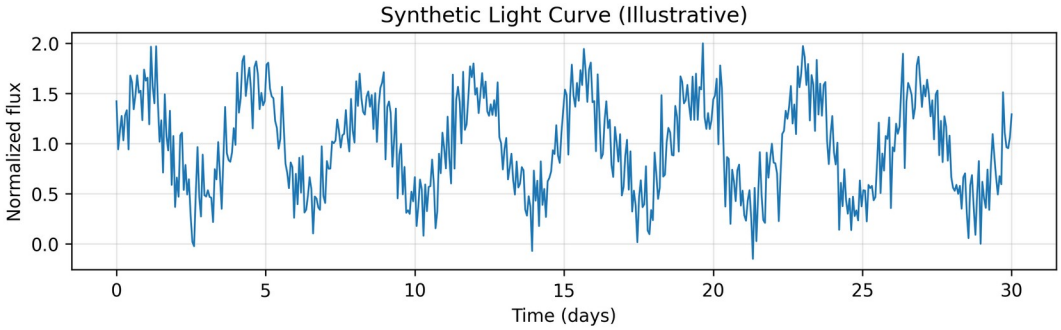

Figure 4. Synthetic light curve (illustrative).

Modern astrophysics is computationally intensive: simulations of galaxy formation, radiative transfer, and gravitational dynamics require high-performance computing (HPC), while observational surveys require pipelines for calibration, catalog construction, and statistical inference. This matters here because it clarifies what ‘using’ a source means in astrophysics: astrophysics uses numerical models, data pipelines, and reproducible computation. The relevant contribution from Sanskrit astronomical traditions is not physical equivalence but historical algorithm design—an early example of building compute systems under constraints.

7.6 Time standards, coordinate frames, and engineering details

A calendar engine that aims for scientific reproducibility must explicitly specify time standards and coordinate frames. Premodern texts assume local observational contexts; modern computation must handle UTC offsets, leap seconds, and differences between civil time (UTC) and dynamical time (TT). Similarly, longitude and latitude conventions must be explicit, along with the definition of ‘sunrise’ (geometric horizon, refraction model, elevation). These details often dominate discrepancies between implementations.

7.7 Software architecture for a calendrical rule engine

A robust rule engine can be organized as follows:

- Input module: ephemeris provider, coordinate frames, time standards

- Core computation: event detection for tithi/nakshatra/yoga/karana
- Rule layer: text-derived predicates for festival and ritual timing
- Validation suite: unit tests and benchmark dates
- Export layer: human-readable panchang outputs, datasets, and visualizations

From a software engineering perspective, this is a clean separation of concerns: numerical astronomy in one layer, textual rules in another, and presentation outputs in a third. It also supports incremental improvement: one can swap ephemeris sources or numerical methods without rewriting the full system.

8. Digital Humanities and NLP for Sanskrit Astronomical Corpora

8.1 Why digital methods matter

Sanskrit astronomical texts are scattered across manuscripts and editions, written in diverse scripts and transmitted with variation. Digital methods enable reproducible scholarship: transcriptions can be compared, errors tracked, and analyses rerun. They also make the field more accessible: students and researchers without local access to manuscripts can still work with high-quality digitized sources.

8.2 Manuscript digitization and open repositories

Open digitization initiatives by major libraries provide a foundation for computational research. Digitized images allow citation, sharing, and reprocessing (for example, OCR and layout analysis). The research benefit is immediate: instead of relying on a single printed edition, one can compare multiple witnesses and quantify variation.

8.3 OCR and layout workflows

OCR for Sanskrit faces challenges: ligatures, degraded pages, and nonstandard layouts. Open-source workflows for difficult manuscripts exist, combining layout annotation with text recognition. When fully automatic OCR is unreliable, a semi-automatic workflow—human correction on top of machine suggestions—can still dramatically accelerate corpus construction. In project design, it is important to select targets strategically: printed Devanagari editions may yield high OCR accuracy and can bootstrap the lexicon and domain term list, which can later improve manuscript transcription through language models or constrained decoding.

8.4 Sanskrit NLP: segmentation, morphology, and parsing

Sanskrit processing requires sandhi segmentation and morphological disambiguation. Tooling exists for morphology and segmentation, and neural toolkits have been proposed for tasks including segmentation, morphological tagging, and dependency parsing. For astronomy texts, domain adaptation may be needed because technical vocabulary differs from epic or philosophical corpora.

8.5 Corpora and evaluation

Linguistically annotated corpora enable evaluation. A corpus with sandhi splits and morphological tags can serve as a benchmark for pipeline accuracy, while astronomy-specific annotations (planets, units, procedures) enable downstream extraction of rules. Evaluation should report both overall accuracy and error types, since sandhi and technical abbreviations produce characteristic mistakes.

8.6 Linking text, rules, and computed outputs

The central digital-humanities innovation in this project is to link textual passages to formal computational objects. A verse is not only quoted; it is mapped to a procedure, implemented, and validated. This makes interpretation accountable: if one claims that a passage specifies a computation, the resulting output can be examined and compared.

8.7 A worked example workflow (conceptual)

A worked example illustrates the workflow. Suppose a passage describes how to determine a particular tithi-based observance. The pipeline would: (i) identify the passage; (ii) segment and lemmatize it; (iii) tag domain terms (tithi, sunrise, month name); (iv) translate the rule into an event-detection specification; (v) run the computation for a chosen location; and (vi) compare the resulting timing to a reference calendar. The research value is that both the textual interpretation and the numerical output are documented and reproducible.

9. Research Design: A Reproducible Computational Project

9.1 Data and sources

Select a small but representative corpus: a set of Vedic/Upanishadic passages about time and cosmology; Vedanga Jyotisha sections about calendrical cycles; and a technical subset from siddhanta materials focused on solar and lunar motion. Use digitized sources where licensing permits, and store metadata (repository, shelfmark, page/folio identifiers) alongside transcriptions.

9.2 Pipeline overview

- Digitization: acquire images and produce initial transcriptions (OCR where feasible).

- Normalization: consistent transliteration and Unicode normalization.
- Segmentation and morphology: sandhi splitting, lemma assignment, POS tagging.
- Domain annotation: tag astronomical concepts (Sun, Moon, planets), time units, and procedural statements.
- Rule extraction: translate procedural text into formal predicates and equations.
- Implementation: build a rule engine that computes calendar elements and event times.
- Validation: compare outputs with modern implementations and ephemerides; report error distributions.

9.3 Data model and knowledge graph

Represent the corpus in a knowledge graph that connects TextPassage nodes to AstronomicalConcept nodes and ComputationRule nodes. This supports queries such as: Which rules depend on nakshatra ordering? Which passages specify solstice markers? Which rules produce the largest divergence from modern ephemerides?

9.4 Validation metrics and error analysis

Proposed metrics include:

- Segmentation accuracy (token-level F1) against a gold standard (where available).
- Rule extraction accuracy (human-validated mapping from passage to procedure).
- Calendar accuracy (difference in event times compared to reference ephemerides).
- Robustness (sensitivity to location, time standard, and ephemeris source).
- Explainability: ability to trace an output back to a specific rule and textual passage.

Error analysis should be qualitative as well as quantitative. For example, if an observance time differs by hours between implementations, the cause may be sunrise definition, time standard

conversions, or a text's assumption about local horizon. Documenting these causes strengthens both the computational artifact and the historical interpretation.

9.5 Risk register and mitigation

Common project risks and mitigation strategies:

Risk: Low OCR accuracy on manuscripts. Mitigation: Start with printed editions; use semi-automatic correction; treat manuscripts as a second phase.

Risk: Ambiguity in Sanskrit terms. Mitigation: Preserve alternatives; annotate uncertainty; consult lexicons and secondary literature.

Risk: Scope creep (too many texts). Mitigation: Limit corpus; focus on one calendar family and a small set of rules.

Risk: Validation disputes. Mitigation: Specify reference method; run comparisons across multiple references; report uncertainty.

9.6 Expected outcomes

A strong outcome is not a controversial claim about "ancient astrophysics." It is a dataset, a codebase, and an empirical report: a citable corpus of passages, an implemented calendrical engine, and quantified comparisons. Such outputs can support further scholarship in history of science, Sanskrit studies, and computational astronomy.

10. Ethics, Epistemics, and Communication

Work that combines sacred texts and modern science is socially sensitive. It is essential to avoid misleading equivalences and to communicate with precision.

10.1 Avoiding pseudoscience while respecting tradition

Respect for tradition does not require scientific overclaim. A paper can honor the cultural and intellectual significance of Vedic literature while clearly stating that modern astrophysics relies on observation, measurement, and mathematical modeling.

10.2 Transparency about uncertainty

Dating of texts, manuscript transmission, and translation choices often involve uncertainty. A research paper should report uncertainty explicitly and avoid presenting interpretive hypotheses as settled fact.

10.3 Reproducibility norms

Because the project includes computation, it should follow reproducibility standards: version-controlled code, environment specifications, test vectors, and clear licensing for any datasets.

10.4 Communication guidelines

Prefer language such as:

- “historically adjacent” rather than “identical”
- “comparative philosophy” rather than “proof”
- “algorithmic reconstruction” rather than “ancient prediction”
- “approximation scheme” rather than “exact modern value”

11. Conclusion

Modern astrophysics does not use the Vedas and Upanishads as scientific data or theory. Yet these texts remain important for understanding early cosmological imagination and for explaining why timekeeping became culturally central. Vedanga Jyotisha provides a bridge from

ritual calendars to rule-based astronomical computation, and later siddhanta traditions developed substantial mathematical astronomy. Computer science offers the most robust contemporary connection: algorithms for calendar computation, reproducible implementations, and digital-humanities methods to digitize and analyze Sanskrit astronomical corpora.

This paper reframes the original question into a credible research agenda: build a corpus; extract computational rules; implement them; validate outputs; and interpret cosmological motifs with genre-sensitive caution. The result is both academically defensible and practically valuable—producing datasets, tools, and analyses that can support future interdisciplinary scholarship.

References

Aghanim, N., et al. (2020). Planck 2018 results. I. Overview and the cosmological legacy of Planck. Astronomy & Astrophysics, 641, A1.

Bodleian Libraries. (n.d.). Sanskrit and South Asian Manuscripts Digitization Project. University of Oxford.

Cambridge Digital Library. (n.d.). Sanskrit Manuscripts collection. University of Cambridge.

Dershowitz, N., & Reingold, E. (2018). Calendrical Calculations: The Ultimate Edition. Cambridge University Press.

Dershowitz, N., & Reingold, E. (n.d.). Indian Calendrical Calculations (technical paper accompanying calendrical algorithms).

European Space Agency. (n.d.). Planck and the cosmic microwave background. ESA Science & Exploration.

Jamison, S. W., & Brereton, J. P. (2014). The Rigveda: The Earliest Religious Poetry of India (3 vols.). Oxford University Press.

NASA. (n.d.). Big Bang and the evolution of the Universe. NASA Astrophysics: Physics of the Cosmos.

NASA. (n.d.). WMAP overview. NASA Science.

Olivelle, P. (1998). The Early Upanishads: Annotated Text and Translation. Oxford University Press.

Pingree, D. (1973). The Mesopotamian Origin of Early Indian Mathematical Astronomy. Journal for the History of Astronomy, 4, 1–12.

Plofker, K. (2009). Mathematics in India. Princeton University Press.

Sanskrit Heritage Site (INRIA). (n.d.). Sanskrit processing tools: dictionary, morphology, segmentation, parsing.

Sandhan, J., et al. (2023). SanskritShala: A Neural Sanskrit NLP Toolkit with Web-Based Interface. ACL System Demonstrations.

Subbarayappa, B. V., & Sarma, K. V. (1985). Indian Astronomy: A Sourcebook. Nehru Centre / INSA.

Appendix A: Knowledge Graph Schema for Calendrical Concepts

This appendix sketches a concrete schema for representing calendrical concepts and computational rules extracted from Sanskrit texts.

Entities:

• TextSource {repository, shelfmark, edition, script, license}

• TextPassage {source_id, location, Sanskrit_text, transliteration, translation, notes}

• AstronomicalConcept {term, lemma, type (object/unit/event), definition, cross_refs}

• ComputationRule {inputs, procedure, assumptions, outputs, citations}

• ValidationResult {rule_id, epoch, location, reference_method, error_metrics}

Relations:

• mentions(TextPassage, AstronomicalConcept)

• defines(TextPassage, ComputationRule)

• depends_on(ComputationRule, AstronomicalConcept)

• produces(ComputationRule, ValidationResult)

A.1 Example concept list

- tithi (lunar day)
- paksha (fortnight)
- amavasya / purnima (new/full moon events)
- nakshatra (lunar mansion)
- saṅkrānti (solar ingress)
- ayana (solstitial half-year)
- muhūrta (time subdivision)

A.2 Example graph queries

- Retrieve all passages that define a rule involving solstice markers, grouped by text.
- List all computation rules that depend on nakshatra ordering and compare their outputs.
- For a chosen epoch, compute tithi transition times across methods and plot error distributions.
- Detect terminological drift: how does a given nakshatra name appear across corpora and scripts?

Appendix B: Pseudocode for Event-Detection in Calendar Computations

The core computational task in many calendar elements is event detection—finding the time at which a function crosses a threshold.

Example: tithi boundary at angle = k × 12 degrees, where angle = (λ_moon − λ_sun) mod 360.

Pseudocode (language-agnostic):

```
function find_boundary(t0, t1, target_angle):
    f(t) = wrap360(lambda_moon(t) - lambda_sun(t)) - target_angle
    while not bracketed(f(t0), f(t1)):
        expand_interval()
    for i in 1..N:
        tm = (t0 + t1) / 2
        if sign(f(t0)) == sign(f(tm)):
            t0 = tm
        else:
            t1 = tm
    return (t0 + t1) / 2
```

Note: In practice, one may use a faster method (e.g., Newton or secant) after bracketing. Bisection is shown because it is robust and easy to validate.

Appendix C: Annotation Guidelines for Sanskrit Astronomy Texts

Annotation layers recommended for the corpus:

- Token layer: surface form plus sandhi-split tokens.
- Morphology layer: POS, lemma, inflectional features, and confidence score.
- Domain entity layer: Sun, Moon, planets, nakshatra names, months, time units, observational anchors (solstice/equinox).
- Procedure layer: tag imperative/procedural spans (“compute,” “add,” “divide,” “when X then Y”).
- Uncertainty layer: multiple glosses or interpretations with rationale.

Quality control measures: inter-annotator agreement, spot checks against morphology tools, and a transparent changelog for normalization decisions.